PB88-103007

Regional Field Study Design for Evaluation of Eulerian Acidic Deposition Models

(U.S.) Environmental Protection Agency
Research Triangle Park, NC

Sep 87

U.S. Department of Commerce
National Technical Information Service
NTIS

PB88-103007

EPA/600/D-87/268
September 1987

REGIONAL FIELD STUDY DESIGN FOR EVALUATION OF EULERIAN ACIDIC DEPOSITION MODELS

by

H. Michael Barnes
U.S. Environmental Protection Agency
Research Triangle Park, North Carolina 27711

D. Alan Hansen
Electric Power Research Institute
P.O. Box 10412
Palo Alto, California 94303

Maris Lusis
Ministry of the Environment
880 Bay Street
Rexdale, Ontario M5S 1Z8, Canada

ATMOSPHERIC SCIENCES RESEARCH LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
RESEARCH TRIANGLE PARK, NORTH CAROLINA 27711

REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL
INFORMATION SERVICE
SPRINGFIELD, VA 22161

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA/600/D-87/268	2.
3. RECIPIENT'S ACCESSION NO. PB88-103007	
4. TITLE AND SUBTITLE REGIONAL FIELD STUDY DESIGN FOR EVALUATION OF EULERIAN ACIDIC DEPOSITION MODELS	5. REPORT DATE September 1987
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) H. M. Barnes, Jr. et al. (See title page for other authors)	8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Same as Box 12	10. PROGRAM ELEMENT NO. N104/C/05/05/6058-01 FY-87
	11. CONTRACT/GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS Atmospheric Sciences Research Laboratory - RTP, NC Office of Research and Development U.S. Environmental Protection Agency Research Triangle Park, North Carolina 27711	13. TYPE OF REPORT AND PERIOD COVERED
	14. SPONSORING AGENCY CODE EPA/600/09

15. SUPPLEMENTARY NOTES

16. ABSTRACT

The U.S. Environmental Protection Agency and the Electric Power Research Institute, the Ontario Ministry of the Environment, and the Atmospheric Environmental Services have funded the development of two complex regional scale Eulerian acid deposition models, RADM and ADOM. There is a consensus among the scientific community that the models should be evaluated as to their skill in predicting output parameters and to establish their credibility. This paper describes the measurements planned in a large field study to provide data for this evaluation. The plan is divided into two parts: operational and diagnostic. The operational evaluation assesses the models' ability to predict air concentrations and deposition patterns. The diagnostic evaluation is an assessment of the behavior of the model and its individual components, transport, dry deposition, gas phase chemistry, emission inventory, etc. Plans for an extensive network of ground sampling stations and intensive aircraft measurements over a two-year period are discussed.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT RELEASE TO PUBLIC	19. SECURITY CLASS *(This Report)* UNCLASSIFIED	21. NO. OF PAGES 14
	20. SECURITY CLASS *(This page)* UNCLASSIFIED	22. PRICE A02 9.95

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE

NOTICE

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

REGIONAL FIELD STUDY DESIGN FOR EVALUATION OF EULERIAN ACIDIC DEPOSITION MODELS

by

H. Michael Barnes
U.S. Environmental Protection Agency
Research Triangle Park, North Carolina 27711

D. Alan Hansen
Electric Power Research Institute
P.O. Box 10412
Palo Alto, California 94303

Maris Lusis
Ministry of the Environment
880 Bay Street
Rexdale, Ontario M5S 1Z8, Canada

INTRODUCTION

During meetings held in 1980 and 1981 under the Memorandum of Intent on Transboundary Air Pollution the need for regional Eulerian acid deposition models became apparent. The models available then, Lagrangian transport and deposition models, simplify the known science especially concerning oxidant formation. The conclusion was that the complex transformation processes leading to oxidant and acid formation required an Eulerian framework.

Two Eulerian models are being developed presently. The U.S. Environmental Protection Agency (EPA) is supporting the development of the Regional Acid Deposition Model (RADM) at the National Center for Atmospheric Research (NCAR). The Ontario Ministry of the Environment (OME), the Atmospheric Environment Services (AES), and the Electric Power Research Institute (EPRI), and the Umweltbundesamt (West Germany) are supporting the development of the Acid Deposition and Oxidant Model (ADOM). Both models are state-of-the-science, specifically designed for application to the complex questions implied by the policy issues.

The two models are similar in nature and complexity. Both contain meteorological drivers, gas-phase and aqueous-phase chemistry components, precipitation (wet deposition) and dry-deposition modules. Both require gridded emission inventories, upwind concentrations, pollutant concentrations, surface characteristics, etc., as inputs; the outputs of the models (layer averaged concentrations, chemical conversion rates, wet deposition fluxes, surface fluxes, etc.) are likewise similar.

Although these two models are the most comprehensive state-of-the-science Eulerian models available, there is a consensus of the need for a model evaluation. The objectives of the evaluation are to establish the skill of these Eulerian models and to assess and establish their credibility. Although the RADM and ADOM have been shown to be capable of addressing policy issues, their predictions will not gain wide acceptance unless the models have been subjected to a credible model evaluation program. An important component of the modeling program must be the establishment of credibility by comparisons with observations. Unless the models are evaluated, the unquantified uncertainties associated with model predictions will severely limit the use of the models in making policy decisions.

EXPERIMENTAL

EPRI[1] has divided model evaluations into two types: (1) "Operational evaluation" assesses the models' ability to predict deposition and air concentrations patterns for effects applications and policy analysis. (2) "Diagnostic evaluation" is an assessment of the behavior of the model and its individual modules. The former evaluation will require surface measurements that produce seasonal and annual averages over Eastern North America; the latter evaluation requires measurements during episodes (1-6 days) to determine whether the models simulate crucial features of transport, transformation, and deposition. A successful diagnostic evaluation would establish confidence in the ability of the models to simulate other than present conditions.

Daily Surface Monitoring Network

The monitoring network is essential for the operational evaluation of the models. The surface network will be composed of two existing and two planned networks sponsored by different organizations (see Figure 1). The existing networks are the OME-APIOS and the AES-CAPMoN, both in existence in Canada for several years. The new networks to be implemented in 1988 are the EPA ME-35 and the EPRI-OEN (Operational Evaluation Network). Each network should follow a compatible protocol for measuring deposition, air quality and meteorology. The surface network will eventually be of national scale, but the initial primary focus will be in the eastern half of the United States and Southeastern Canada.

In Table I we list the variables specified to be monitored in the surface network for the model evaluation. In the particular case of the EPA the ME-35 network will be collocated with existing NTN-National Dry Deposition Network (NDDN) sites which will be in operation later this year. The National Trends Network (NTN) will eventually comprise 150 sites operated by the U.S. Geological Survey (USGS) for wet deposition monitoring. The NDDN will comprise 100 air quality monitoring sites operated for the EPA by its Environmental Monitoring and Systems Laboratory (EMSL/RTP). The first 35 NDDN sites have been selected with consideration of the other networks to optimize the capability of the combined networks to resolve spatial patterns and achieve acceptable levels of uncertainty.

TABLE I. Monitoring variables for surface network

Gases (24 hour average):	SO_2, HNO_3, NH_3, NO_2*,
Gases (1-hour average):	O_3**, HOOH*, PAN**. NO_2**
Particles (24 hour average):	<~10μm, NH_4^+, SO_4^{2-}, NO_3^- and Cl^-
Precipitation (24-hour average):	H^+ (as pH), conductivity, SO_4^{2-}, NO_3^-, NH_4^+, Cl^-, Na^+, K^+, Ca^{2+}, and Mg^{2+}
Meteorology (3 hour average):	Surface wind speed and direction, temperature, pressure*, precipitation, relative humidity and insolation

*EPA only **EPRI only

The EPA and its Atmospheric Sciences Research Laboratory (ASRL) will be responsible for operating the ME-35 in such a way as not to perturb the week-long sampling protocol of the collocated NDDN and NTN-Wet. The NDDN will measure and record continuously the meteorological variables and O_3. Hourly averages will be computed from these data and made available for model

evaluation with the ME-35 and OEN data. The ME-35 will operate separate gas, particle, and precipitation samplers.

The OEN sites operated by the EPRI are shown on Figure 1. Wherever possible, present UAPSP sites shall be used for the expanded monitoring program to maintain data continuity in the existing OEN program. The species listed in Table I will be monitored (except for HOOH) in the OEN; additionally, gas-phase PAN (hourly) and particulate mass (24 hour) will be determined. The EPA and the EPRI networks will use the same methodology for the precipitation samples.

The contributing sites in Canada will all contain precipitation sampling equipment, the samples being analyzed for the species shown in Table I (except for conductivity). At many of the sites, air quality will also be monitored, using filter pack equipment, with analysis for particulate sulfate, nitrate and gaseous nitric acid, and SO_2 (note: no particle sizing is expected at these sites). A few sites will make measurements of PAN, NH_3, NO_2, and HOOH. Collocated ozone monitors will supplement air data at several sites, although most zone data in this area will result from a dense network of ozone monitors in southern and central Ontario. On-site meteorological data will generally be unavailable. Most of the wet data will come from nearby weather stations.

In addition to the operational evaluation and the resulting ground network to provide data, it is proposed to have field projects that relate to diagnostic evaluations of the models' inputs and their components. The EPA in cooperation with other interested parties will be involved in the conduct of the planned tasks. The individual projects and potential sponsors are listed in Table II.

TABLE II. DIAGNOSTIC EVALUATION ACTIVITIES

	Project Title	Sponsor
1.	Conduct Vertical Profiles Over Modeling Domain	EPA, AES
2.	Determine Subgrid Deposition Variability	EPA
3.	Upgrade 1985 NAPAP Emission Inventory	EPA, EPRI, OME, AES
4.	Evaluate Inflow Boundary Conditions	NOAA, OME, AES
5.	Evaluate Wet Deposition Module	DOE, EPRI
6.	Evaluate Dry Deposition Module	EPA, DOE, AES
7.	Evaluate Atmospheric Transport Module	NOAA, DOE, EPRI
8.	Evaluate Gas-Phase Chemistry Module	NOAA, EPRI, OME, AES
9.	Comparison and Performance Evaluation	EPA, EPRI, OME, AES

We will now give a brief description of each of the projects.

Vertical Profiles. A comprehensive evaluation of the RADM and ADOM models will require comparisons against the three-dimensional fields of species concentrations averaged over several layers within selected grid squares and the flux of those species through vertical and across cell boundaries. Although plans are not finalized, the use of at least four research-grade aircraft flying 600 km segments including vertical spirals up to 4 km are envisioned for the two-year study period.

Sub-Grid Deposition Variability. The purpose is to deploy 5-15 additional sampling sites within selected modeling domain grid cells to provide estimates of air concentration and wet deposition variability within a cell. The VAR stations (see Figure 1 for possible locations) are to be simpler but compatible with the ME-35 stations; they will consist of 24-hour average air concentration samplers and 24-hour average precipitation collectors.

1985 NAPAP Annual Emission Inventory. The emission inventory to be used in the evaluation may be deficient in cases for which individual episodes (2-6 days) are examined. Proposed work for this project includes: (a) survey of the 100 largest SO_2 and NO_x sources during intensive study periods; and (b) limited field measurements to estimate NO_x and volatile organic compounds (VOC) emission fluxes from selected area sources.

Inflow Boundary Conditions. This project will allow for an evaluation of the input conditions chosen for the models. Aircraft measurements should be made of layered averaged species concentrations and horizontal wind fields along the inflow boundary of the modeling domain.

Wet Deposition Module. The RADM (80 km x 80 km) and ADOM (127 km x 127 km) models employ rather large grid cells with relatively loose vertical resolution. This property prevents them from capturing detailed structure of many storm systems. The experimental approach here would be field measurements that would be made over areas bounded by a small number of adjacent grids; the spatial (10 km) and temporal (equivalent to 0.1 cm precipitation) density would be sufficient to document subgrid storm structure.

Dry Deposition Models. The field study for this project would be conducted in a limited number (10 or less) of areas distributed throughout the model domain to sample the major land use types. In each area a cluster of 4 to 6 sites instrumented to use the inferential technique for estimating dry deposition would be centered around a research grade site. The clusters would encompass 2 to 4 grid cells and the sites could be the same as those discussed above in the Sub-Grid Deposition Variability project.

Atmospheric Transport Module. The output of this project would be a data base of air parcel trajectory information sufficient to evaluate two aspects of the transport module: long-range horizontal transport and vertical transport through storm systems. The Field Study Design Workshop[2] provided several recommendations for experiments: (1) supplement the planned Across North America Tracer Experiment (ANATEX); (2) a smaller-scale tracer experiment located in the modeling domain during one of the intensive sampling periods.

Gas-Phase Models. Since there is no experimental approach known that can isolate all major sources of error arising from the condensation and consolidation of chemical reactions, the recommendation is made that the chemistry modules be evaluated by a combination of explicit theoretical predictions, smog chamber experiments and atmospheric observations of concentrations and ratios of reactive intermediates (VOCs) and products (e.g., HOOH, CH_2O, PAN, NO_3, etc.).

RESULTS

Although the comparison and Performance Evaluation project is listed in Table II, we discuss it in this section for it is ultimately the results of the evaluation of the field study data base that we seek.

No final decisions have been reached on methods that should be applied in making the comparisons or on the criteria to be used in judging the model performance. Specifications will be developed in an EPA workshop scheduled for July 1987. It is possible to discuss likely approaches that will be included in the final evaluation protocol, although this discussion is not intended to be complete. The anticipated approaches to be used are: (1) Conventional Point-to-Grid Comparison, and (2) Pattern Comparisons.

Point-to-grid comparison for evaluation is based on the suggestions of the American Meteorological Society Workshop on Dispersion Model Performance[3] (Fox). This approach has been used previously to evaluate regional scale air quality models[4] (Ruff) in a project sponsored by EPRI and in the evaluation work of the Memorandum of Intent[5]. The difference, d_i, between an observed variable, C_{oi}, and the modeled value, C_{mi} is the basic quantity used in assessing the model performance. A "perfect" model, driven by perfect initial and boundary conditions, would yield $d_i = 0$ at every comparison point. This level of agreement is not possible due to the compromises made in developing the models and the stochastic component of many of the modeled variables. Given the expected problems of comparing individual point values with gridded model averages, EPRI and the MOI projects used the following criteria for model "perfection":

1. $\langle d \rangle = 0$ where d is the bias
2. σ_d, the root mean square error, is small compared to the standard deviation of C_m, σ_{cm}
3. d_i is not a function of the model input parameters.

In addition, the EPRI study prepared scatterplots of C_{oi} versus C_{mi}, showing a perfect line fit, the correlation coefficient, and the mean bias, $\langle d \rangle$.

The principal advantage of this statistical approach to evaluation lies in its conventionality. If we assume that the differences, d_i, are normally distributed, then the statistical tests can be applied to the data and interpretations of the results are relatively straightforward. A major disadvantage of this technique is the basic problem of comparing a value at one point within a grid cell (80 km x 80 km in the RADM case) with the model averaged

value. The representativeness of the single point value is always open to question. The problem has been approached[4] by selecting regionally representative sites that were identified by autocorrelation analysis on hourly averaged concentration data. The hypothesis was that high correlations between one- hour and three-hour measurements indicated regional sites and low correlations denoted sites influenced by local sources on shorter time scales.

An additional problem with this technique is that the statistical measures are sensitive to spatial displacements between observed and modeled fields. This is an especial problem for areas with large gradients. In such instances, large differences (d_i) can result from slight discrepancies in the shape and placement of the modeled and observed fields. The problem can be addressed by subjectively comparing the analyzed and model derived fields, but this approach suffers from the inherent shortcomings of subjective judgment.

Some of the shortcomings discussed above can be overcome by using pattern comparisons. This technique is also more amenable to the ultimate use of these models, Governmental policy analysis. The EPA is interested in the models' being able to compute seasonal and annual deposition to areas the size of States or portions of States. A method must be available that is capable of converting individual point measurements (such as in-the-ground networks) into estimated fields. One possible technique is kriging[6]. Kriging is an interpolation technique that uses a weighted moving average to estimate functional values between known points. The method can be applied to both observed and predicted data and, if properly applied, may address both types of data for intercomparison purposes.

The basic kriging method (simple kriging) is based on two assumptions: (a) the expected value of a function is constant within the domain of interest; and the difference between data values at two points is a function only of the distance between them (the variance is isotropic). Neither one of these assumptions apply rigorously to deposition distributions in eastern North America.

Kriging has several advantages in situations where the assumptions are valid. One is that it can estimate the variance at each point for which an estimate is made. It can also be used to select optional locations for additional measurements and it can be used to put an error or confidence band on the location of an interpolated isopleth. An additional advantage is that the equations used to derive the optimal weights for the weighted moving average does not depend on the data values. This means that single large values are not overly influential in the final estimates.

Kriging, however, does have distinct limitations. The most obvious ones are the fundamental assumptions of constant expected value and isotropy. If these assumptions are determined to be invalid, then universal kriging might be applied. Universal kriging can be used if the trends in the data are mild enough to be represented by low order polynomials.

A fundamental problem with kriging is in interpreting the meaning of the estimates of kriging error. If the assumptions are not valid for simple or universal kriging, then the validity of the estimates is questionable.

Another pattern comparison technique that can be used is principal component analysis (PCA). It is essentially a technique for multivariate analysis of complex systems characterized by a large number of interdependent variables. By mathematical analysis, this large number is reduced to a smaller number of independent, uncorrelated variables. The ordering or ranking of the components (first, second, third, etc.) allows one to account for linear combinations of variables that explain the most, second most, etc., of variability in the data. The principal attractiveness of the technique is in the significance of the ordering of the components. In the Eulerian model evaluation, PCA would be applied to observed and modeled data at a series of points. If the model is an accurate representation of the system, then the major components should be the same as those established from the measurements.

Several disadvantages exist with the PCA technique. A major concern exists about the uncertainty about the robustness of individual components (the technique is sensitive to outliers). Also, PCA places severe demands on the data set necessary; more demands than kriging does. PCA requires simultaneous time resolved measurements of each variable for a complete data analysis to be done. In terms of data completeness, PCA requires at least a 90% valid data base.

CONCLUSIONS

The necessity for a field study to obtain observational data to be used in evaluating the RADM and ADOM has been determined by the clients, EPA, EPRI, OME, and AES. The models must demonstrate the capability of simulating seasonal and annual ambient air quality and deposition over the northeastern United States and southern Canada within some definable error. In addition, the models must simulate the individual processes of transport, transformation, scavenging and deposition sufficiently correctly that the calculated seasonal and annual averages are the result of correct science.

The plan for the field study and subsequent evaluation presented here is the result of a cooperative effort over the past two years by the clients. Various workshops were sponsored to produce specifications for the surface network, the overall field study, model evaluation protocols and Quality Assurance protocols. These workshops have resulted in the plan summarized in this article. The plans discussed are the best suggestions of a consensus group. Funding availability during the next two years will determine which of the projects discussed here actually occur. Thus far, the funds for the ground-surface network, the vertical profile project, the EPA subgrid deposition study and the model evaluation have been committed by the various agencies to perform the tasks described here. Suggested sponsors for the other projects have been made in Table II.

Because a study of this magnitude has not been performed in the past, a great deal of coordination among the agencies is required and is being implemented as evidenced by the workshops held by the clients these past two years. The study will be non-traditional in nature because of its scope and magnitude and the model evaluation as described in this paper reflects that non-traditional approach.

REFERENCES

1. S. D. Reynolds, C. Seigneur, T. E. Stoeckenius, G. E. Moor, R. G. Johnson, and R. J. Londergan, Operational Validation of Gaussian Plume Models at a Plains Site, EA-3076, Electric Power Research Institute.

2. Ibid.

3. D. G. Fox, "Judging Air Quality Model Performance," Bull. Amer. Meteorol. Soc., 62, No. 5, 599 (1981).

4. R. E. Ruff, K. C. Nitz, F. L. Ludwig, and C. M. Bhumralkar, Regional Air Quality Model Assessment and Evaluation, EA-3671, Electric Power Research Institute.

5. United States-Canada Memorandum of Intent (MOI) Work Group 2, Final Report on Atmospheric Sciences and Analysis, Washington, D.C. and Toronto, Canada (1982).

6. M. G. Barnes, "The Use of Kriging for Estimating the Spatial Distribution of Radionuclides and Other Spatial Phenomena," TRANS-STAT: Statistics for Environmental Studies, No. 13, PNL-SA-9051, Pacific Northwest Laboratory, P.O. Box 999, Richland, Washington (1980).

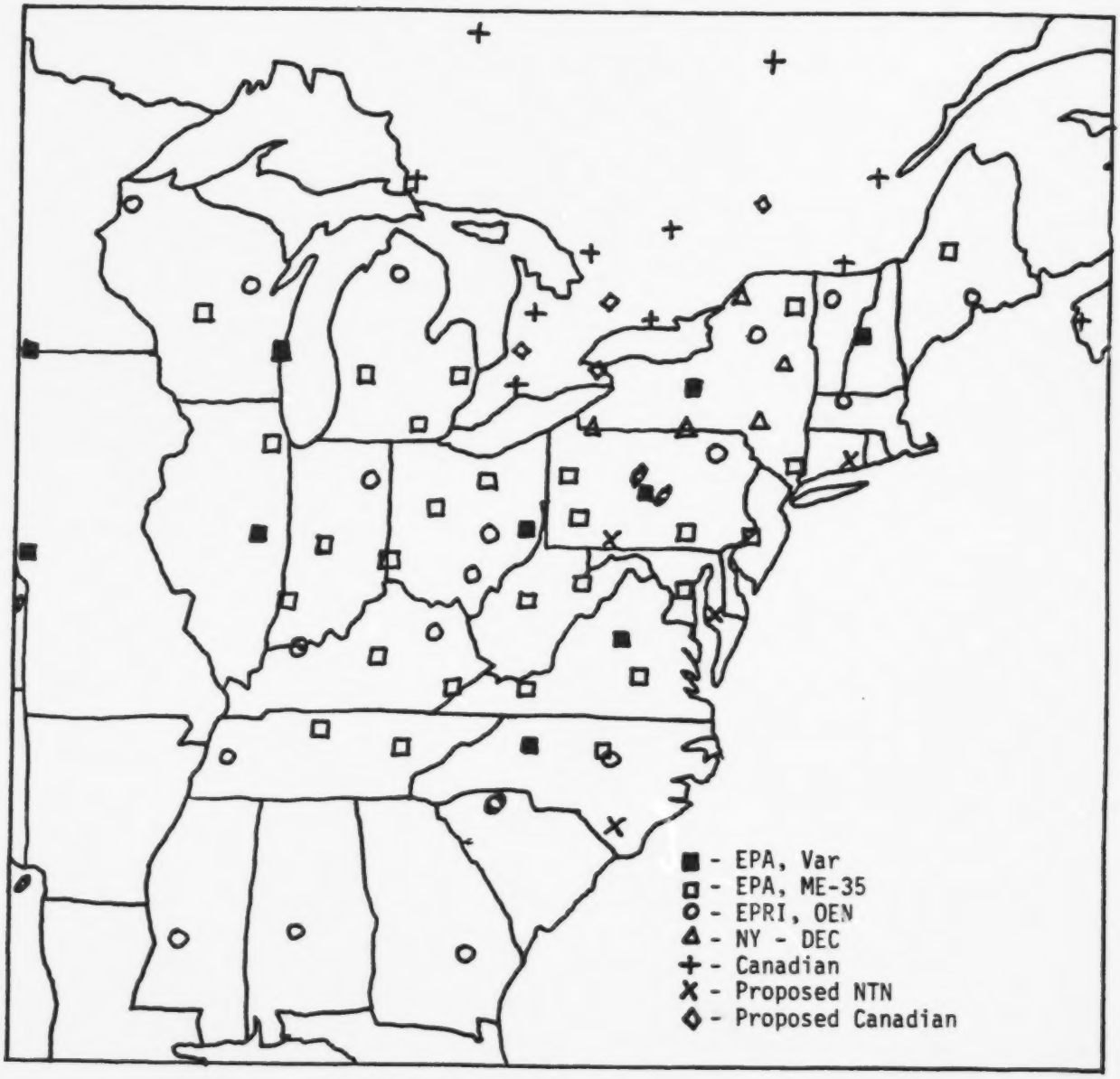

FIGURE 1. SURFACE NETWORK SITES